Nan Rubin, WNET, Visits the Library

By Staff of the National Digital Information Infrastructure and Preservation Program

What happens to the programs produced by public television once they have aired? Today, most of these programs are produced in digital formats, and they are at risk of being lost if they are not preserved. The issue is part of the overall problem that libraries and archives of all types are struggling with: How to ensure the availability of materials that are "born digital" for future generations.

The Library's digital preservation program, called the National Digital Information Infrastructure and Preservation Program (NDIIPP), has partnered with many institutions and organizations as part of its plan to construct a "digital preservation network" of institutions committed to collecting and preserving digital information.

Nan Rubin, from public television station WNET ("Thirteen"), New York, visited the Library on March 8 to discuss the NDIIPP-funded Preserving Digital Public Television partnership project. WNET is the lead partner in the project, which also includes television station WGBH, Boston; program distributor the Public Broadcasting Service (PBS); and New York University (NYU). WNET and WGBH produce more than 60 percent of the program content aired on public television.

The project is designing an archive for long-term preservation of public television programs being produced in a "born-digital" environment. Rubin began her talk by showing a PBS video montage of the diverse materials produced by public television, including snippets of "Sesame Street," Julia Child's pioneering cooking programs, and a number of documentary and news programs. The video provided a vivid example of the rich cultural heritage that may be endangered unless steps are taken to ensure its preservation.

Rubin noted that almost all current WNET productions are shot and edited digitally, and that broadcasting itself is virtually all-digital these days. In spite of this, the public television community has been able to apply only limited resources to preserve the rapidly expanding quantity of digital materials being produced.

"We've got many different acquisition formats, and they're changing all the time," said Rubin. "If you were to list every type of digital media that has been used over the last 20 years, we probably have to deal with it." In some cases, the digital media are conventional videotapes. In more recent cases, however, new content takes "media-less" form, that is, as digital files maintained in server systems.

The video production process was never simple, even when based on conventional videotapes, whether analog or digital. Media-less digital production, however, has introduced a number of complexities into the equation. Digital materials enter the production workflow from a variety of sources, often through the work of independent

producers. These producers do their work on widely divergent types of equipment of variable age, make, capability, and quality. These varying acquisition formats can be normalized to a degree when they are imported for editing into the centralized production systems maintained by the stations, but the process still leaves room for variation in the format and quality of the production materials.

"We're not in a position to mandate how the programs are shot," Rubin said. "We can mandate the format they come out of the editor, but we can't mandate the format coming into the production stream." The original footage has intrinsic value, since it can be repurposed in a number of different ways, for example, in new programs that can benefit from the same video content. But some materials still end up on the cutting room floor. "In some of our discussion with producers, they said they are shooting 100:1," noted Rubin, meaning that only 1 percent of the total material shot by a camera ended up in the production. Yet not every minute of footage can be preserved. Rubin noted that with "the explosion of digital formats and the rapid reduction in the cost, selection issues [deciding which digital materials should be preserved] become very important."

Regarding the repurposing of finished programs and program segments, Rubin said, "Our broadcast signals travel many different directions inside the station. They go to many different uses, not solely on the air: To the Internet, to DVDs, to close-captioning. They do lots of different things with the technology that are not limited to the things that go out over the air."

Rubin explained that a high-resolution (albeit digitally compressed) version of the finished programs was sent by stations like WNET and WGBH to PBS for use in distribution. This high-resolution version is suitable for long-term preservation, she said. PBS, in turn, creates a more heavily compressed version of each show to be sent to its 355 member stations -- the version that is actually transmitted to home televisions. This version, Rubin observed, would be less well-suited for long-term preservation.

The Preserving Digital Public Television project with NDIIPP has made excellent progress during its first year. "We've started looking at at-risk digital materials," continued Rubin. "With many items, we don't even have the machines available to play them back [because they were produced in currently obsolete formats]. We're looking at the production workflow and actual technical questions, including file formats and metadata wrappers [metadata is information attached to a digital file that helps describe it, much the same way that cataloging data describes a book]. We're in pretty close touch with a number of engineers at PBS to see how they are using formats.

"We're also planning a preservation repository. This means that we want to adopt technical and operational standards so that we can implement something that has longevity that we can maintain over time." One of the key contributions from NYU is the development of a model digital repository to demonstrate several of the project's technical findings.

In the area of preservation metadata, the project has been working to continue the development of the PBCore metadata set, "specifically so we can put some preservation fields in the Public TV metadata," noted Rubin. They've also been reviewing best practices and convening focus groups to provide insight on their potential approaches.

"We felt that we needed to approach user focus groups apart from librarians," she said. "We organized five different focus groups; K-12 teachers; social science historians from higher education; television producers; media studies instructors; and the critics and journalists who write about television. We showed them outtakes from productions and asked them how we should handle the materials. They were all mindful that we should preserve the materials, but we were surprised that the producers wanted public television to provide curatorial support and to take the curatorial decisions out of their hands. The critics were shocked that we were even asking them, as their work is largely focused on what's happening today, but they had interesting comments to add.

"We also have a lot of questions to answer regarding access to the material," she continued. "We've been working with the Library's Section 108 working group to some extent. Public television has never really considered whether it was a 'library' or an 'archive' in section 108 terms, but we're trying to inject ourselves into that arena." [The Section 108 Study Group is a select committee of copyright experts, convened by the Library of Congress, and charged with updating for the digital world the Copyright Act's balance between the rights of creators and copyright owners and the needs of libraries and archives.]

Rubin also noted that the public television project shares a number of areas of common interest with the Motion Picture, Broadcasting, and Recorded Sound Division (M/B/RS). For many years, M/B/RS has worked with PBS to provide a permanent archive for public television content, thus far in analog (physical) form. A second area of common interest concerns M/B/RS planning related to the National Audio-Visual Conservation Center being established in Culpeper, Va., a facility that will maintain extensive digital activities.

"We are eager to begin working directly with the Library on the NAVCC," Rubin stated. "We want to ensure that whatever plans we come up with, or whatever files we have ready to transfer, the Library will be ready to receive them. We're getting in on the ground floor as the NAVCC is being built, and it's nice to get in on the ground floor of that development." Rubin hopes the early results will provide a catalyst to increase awareness for both public television management and viewers.

Musing about the fundraising that is a familiar part of public television's presence for viewers, Rubin asked, "Have you ever heard that your money is going to go to help preserve the materials?" You have not, she said, "because the stations need every dime to support ongoing programs, and preservation and archiving are expensive. After the broadcast goes over the air, the local stations don't feel they have the responsibility. There has never been any money allocated by PBS for preservation purposes. They would like to take some responsibility, but they have no money. …We hope that we'll be

able to come up with some models that will support us and the interests of the American people with the support of the Library."

A Webcast of the event is available at http://www.loc.gov/today/cyberlc. Click on the "Science, Technology" link for the "Preserving Public Television" program.